

# Project 5 - Sentiment Analysis

## 1. Sentiment Analysis on IMDB Reviews dataset

The IMDB Movie Reviews dataset is a **binary sentiment analysis dataset consisting of 50,000 reviews from the Internet Movie Database (IMDb) labeled as positive or negative**. The dataset contains an even number of positive and negative reviews. Only highly polarizing reviews are considered. In this project, you have to predict the number of positive and negative reviews based on sentiments using different classification models, including Logistic Regression, Naïve Bayes, Decision Tree (Random Forest, Adamboost), SVM, and Deep Learning using keras library.

The basic code is given in the IMDB case study folder, including the dataset and application of several ML classifiers. You need to modify the existing code.

The deliverable of this project is a report in latex and pdf, similar to the sample included in the file.

## 2. Sentiment analysis of Airbnb reviews for Greek regions

Airbnb has successfully disrupted the traditional hospitality industry as more and more travelers decide to use Airbnb as their primary accommodation provider. Since its inception in 2008, Airbnb has seen enormous growth, with the number of rentals listed on its website growing each year exponentially. The two theses included in the file study this industry for the Athens region.

In this project, you will study what visitors like and dislike in the region of Crete.

To find out, you will process the reviews to find out what people's likes and dislikes are using Natural Language Processing (NLP) and specifically Sentiment Analysis.

The file contains codes for similar studies in various regions, including Athens.

Other similar case studies can be found on [kaggle.com](https://www.kaggle.com)

The data for the region of Crete can be found at <http://insideairbnb.com/get-the-data/>

You will apply similar classifiers as in the case of the IMDB study

The deliverable of this project is a report similar to the sample included in the project file using latex and pdf.